Exhibit D

Part 6

DETAILED DESCRIPTION OF THE ... the read-back waveform, sampled ... of one sample

circuit **32**. A noise statistics tracker circuit **34** uses the delayed samples and detector decisions to update the noise statistics, i.e., to update the noise covariance matrices. A metric computation update circuit **36** uses the updated statistics to calculate the branch metrics needed in the Viterbi-like algorithm. The algorithm does not require

of the detector circuit **26** of FIG. 1. The detector circuit **28** has a feedback circuit **32** which feeds back into a Viterbi-like detector **30**. The outputs of the detector **30** are decisions and delayed signal samples, which are used by the feedback circuit **32**. A noise statistics tracker circuit **34** uses the delayed samples and detector decisions to update the noise statistics, i.e., to update the noise covariance matrices. A metric computation update circuit **36** uses the updated statistics to calculate the branch metrics needed in the Viterbi-like algorithm. The algorithm does not require replacing current detectors. It simply adds two new blocks in the feedback loop to adaptively estimate the branch metrics used in the Viterbi-like detector **30**.

The Viterbi-like detector **30** typically has a delay associated with it. Until the detector circuit **28** is initialized, signals of known values may be input and delayed signals are not output until the detector circuit **28** is initialized. In other types of detectors, the detector may be initialized by having the necessary values set.

The correlation-sensitive maximum likelihood sequence detector (CS-MLSD) **28** is described hereinbelow. Assume that N>1 channel bits (symbols), $a_1, a_2, \ldots, a_N$, are written on a magnetic medium. The symbols a_i, $i=1, \ldots, N$, are drawn from an alphabet of four symbols, $a_i \in \{+, \oplus, -, \ominus\}$. The symbols '+' and '−' denote a positive and a negative transition, respectively. The symbol '⊕' denotes a written zero (no transition) whose nearest preceding non-zero symbol is a '+' while '⊖' denotes a written zero whose nearest preceding transition is a negative one, i.e., '−'. This notation is used because a simple treatment of transitions as '1's and no transitions as '0's is blind to signal asymmetries (MR head asymmetries and base line drifts), which is inappropriate for the present problem. In FIG. **3** a sample waveform is illustrated. The signal asymmetries and base line shifts are exaggerated in FIG. **3**. FIG. **3** also shows the written symbols $a_1, \ldots, a_{18}$, as well as the samples $r_1, \ldots, r_{18}$ of

By Bayes rule, the joint conditional pdf (likelihood function) is factored into a product of conditional pdfs:

$$f(r_1, \ldots, r_N | a_1, \ldots, a_N) = \prod_{i=1}^{N} f(r_i | r_{i+1}, \ldots, r_N, a_1, \ldots, a_N) \tag{2}$$

To proceed and obtain more concrete results, the nature of the noise and of the intersymbol interference in magnetic recording is exploited.

Finite correlation length. The conditional pdfs in Equation (2) are assumed to be independent of future samples after some length $L \geqq 0$. L is the correlation length of the noise. This independence leads to:

$$f(r_i | r_{i+1}, \ldots, r_N, a_1, \ldots, a_N) = f(r_i | r_{i+1}, \ldots, r_{i+L}, a_1, \ldots, a_N) \tag{3}$$

Finite intersymbol interference. The conditional pdf is assumed to be independent of symbols that are not in the K-neighborhood of $r_i, \ldots, r_{i+L}$. The value of $K \geqq 1$ is determined by the length of the intersymbol interference (ISI). For example, for PR4, K=2, while for EPR4, K=3. $K_l \geqq 0$ is defined as the length of the leading (anticausal) ISI and $K_t \geqq 0$ is defined as the length of the trailing (causal) ISI, such that $K = K_l + K_t + 1$. With this notation the conditional pdf in (3) can be written as:

$$f(r_i | r_{i+1}, \ldots, r_{i+L}, a_1, \ldots, a_N) = f(r_i | r_{i+1}, \ldots, r_{i+L}, a_{i-K_l}, \ldots, a_{i+L+K_t}) \tag{4}$$

Substituting (4) into (2) and applying Bayes rule, the factored form of the likelihood function (conditional pdf) is obtained:

CMU's "i.e." Argument Fails

- Patents use "noise covariance matrices" separately from "noise statistics"

> Because the noise statistics are non-stationary, the noise sensitive branch metrics are adaptively computed by estimating the noise covariance matrices from the read-back data. These covariance matrices are different for each branch of the tree/trellis due to the signal dependent structure of the media noise. Because the channel characteristics in mag-

'839 Patent 2:15-23

> Computing the branch metrics in (10) or (13) requires knowledge of the signal statistics. These statistics are the mean signal values m_i in (12) as well as the covariance matrices C_i in (13). In magnetic recording systems, these

'839 Patent 8:24-27

$$M_i = \log \frac{\det C_i}{\det c_i} + \underline{N}_i^T C_i^{-1} \underline{N}_i - n_i^T c_i^{-1} n_i \quad (13)$$

In the derivations of the branch metrics (8), (10) and (13), no assumptions were made on the exact Viterbi-type architecture, that is, the metrics can be applied to any Viterbi-type algorithm such as PRML, FDTS/DF, RAM-RSE, or, MDFE.

FIG. 3A illustrates a block diagram of a branch metric computation circuit **48** that computes the metric M_i for a branch of a trellis, as in Equation (13). Each branch of the trellis requires a circuit **48** to compute the metric M_i.

A logarithmic circuit **50** computes the first term of the right hand side of (13)

$$\left(\text{i.e. } \log \frac{\det C_i}{\det c_i}\right)$$

and a quadratic circuit **52** computes the second terms of the right hand side of (13) (i.e. $\underline{N}_i^T C_i^{-1} \underline{N}_i - n_i^T c_i^{-1} n_i$). The arrows through the circuits **50** and **52** represent the adaptive nature of the Viterbi-like detector **30**. A sum circuit **53** computes the sum of the outputs of the circuits **50** and **52**.

Where the vector $\underline{w}_i$ is (L+1)-dimensional and is given by:

$$\underline{w}_i^T = [1\ \ w_i(2)\ \ w_i(3)\ \ \ldots\ \ w_i(L+1)]^T \quad (19)$$

$$= \begin{bmatrix} 1 \\ -c_i^{-1}\underline{c}_i \end{bmatrix} \quad (20)$$

Equations (17), (18) and (16) (the circuit **52**) can be implemented as a tapped-delay line as illustrated in FIG. 3B. The circuit **52** has L delay circuits **54**. The tapped-delay line implementation shown in FIGS. 3A and 3B is also referred to as a moving-average, feed-forward, or finite-impulse response filter. The circuit **48** can be implemented using any type of filter as appropriate.

The adaptation of the vector of weights $\underline{w}_i$ and the quantity σ_i^2 as new decisions are made is essentially an implementation of the recursive least squares algorithm. Alternatively, the adaptation may be made using the least mean squares algorithm.

The quantities m_i that are subtracted from the output of the delay circuits **54** are the target response values, or mean signal values of (12). The arrows across multipliers **56** and across square devices **58** indicate the adaptive nature, i.e., the data dependent nature, of the circuit **52**. The weights $\underline{w}_i$ and the value σ_i^2 can be adapted using three methods. First, $\underline{w}_i$ and σ_i^2 can be obtained directly from Equations (20) and (16), respectively, once an estimate of the signal-dependent covariance matrix C_i is available. Second, $\underline{w}_i$ and σ_i^2 can be calculated by performing a Cholesky factorization on the inverse of the covariance matrix C_i. For example, in the $L_i D_i^{-1} L_i^T$ Cholesky factorization, $\underline{w}_i$ is the first column of the Cholesky factor L_i and σ_i^2 is the first element of the diagonal matrix D_i. Third, $\underline{w}_i$ and σ_i^2 can be computed directly from the data using a recursive least squares-type algorithm. In the first two methods, an estimate of the covariance matrix is obtained by a recursive least squares algorithm.

Computing the branch metrics in (10) or (13) requires

diagonal matrix D_i. Third, $\underline{w}_i$ and σ_i^2 can be computed directly from the data using a recursive least squares-type algorithm. In the first two methods, an estimate of the covariance matrix is obtained by a recursive least squares algorithm.

LMS-class (least mean-squares) algorithm that ensures that the mean of the signal samples is close to these target values. In decision feedback equalization (DFE) based detectors or hybrids between fixed delay tree search and DFE, such as FDTS/DF or MDFE, the target response need not be pre-specified. Instead, the target values are chosen on-the-fly by simultaneously updating the coefficients of the front-end and feed-back equalizers with an LMS-type algorithm.

When there are severe nonlinearities in the system (also referred to as nonlinear distortion or nonlinear ISI), a linear equalizer will generally not be able to place the signal samples right on target. Instead, the means of the signal samples will fall at a different value. For example, in a PR4 system, the response to a sequence of written symbols . . . , −, +, ⊕, . . . might result in mean sample target values . . . , 0, 1, 0.9, . . . , while a sequence of written symbols . . . , +, −, ⊖, . . . might result in a sequence of mean sample

This ratio of determinants is referred to as σ_i^2, i.e.:

$$\sigma_i^2 = \frac{det\ C_i}{det\ c_i} = \alpha_i - \underline{c}_i^T c_i^{-1} \underline{c}_i. \tag{16}$$

$L_i D_i^{-1} L_i^T$ Cholesky factorization, $\underline{w}_i$ is the first column of

Where the vector $\underline{w}_i$ is (L+1)-dimensional and is given by:

$$\underline{w}_i^T = [1 \quad w_i(2) \quad w_i(3) \quad \ldots \quad (w_i(L+1)])^T \tag{19}$$

$$= \begin{bmatrix} 1 \\ -c_i^{-1} \underline{c}_i \end{bmatrix}. \tag{20}$$

Equations (17), (18) and (16) (the circuit **52**) can be implemented as a tapped-delay line as illustrated in FIG. 3B. The circuit **52** has L delay circuits **54**. The tapped-delay line implementation shown in FIGS. 3A and 3B is also referred to as a moving-average, feed-forward, or finite-impulse response filter. The circuit **48** can be implemented using any type of filter as appropriate.

feed-back equalizers with an LMS-type algorithm.

When there are severe nonlinearities in the system (also referred to as nonlinear distortion or nonlinear ISI), a linear equalizer will generally not be able to place the signal samples right on target. Instead, the means of the signal samples will fall at a different value. For example, in a PR4 system, the response to a sequence of written symbols . . . , −,+, ⊖, . . . might result in mean sample target values . . . , 0, 1, 0.9, . . . , while a sequence of written symbols . . . , +, −, ⊖, . . . might result in a sequence of mean sample

CMU's "i.e." Argument Fails

- "noise statistics": used by all branch metrics
- "covariance matrices": used by only "correlation-sensitive branch metric"

Specific expressions for the branch metrics that result under different assumptions on the noise statistics are next considered.

'839 Patent 5:56-58

Variance dependent branch metric.

$$M_i = \log\sigma_i^2 + \frac{N_i^2}{\sigma_i^2} = \log\sigma_i^2 + \frac{(r_i - m_i)^2}{\sigma_i^2} \quad (10)$$

'839 Patent 6:15-34

With this notation, the general correlation-sensitive metric is:

$$M_i = \log \frac{det\ C_i}{det\ c_i} + \underline{N}_i^T C_i^{-1} \underline{N}_i - \underline{n}_i^T c_i^{-1} \underline{n}_i \quad (13)$$

'839 Patent 6:66-7:4

(RLS) algorithm. The RLS computes the next covariance matrix estimate $\hat{C}'(\hat{a})$ as:

$$\hat{C}'(\hat{a})=\beta(t)\hat{C}(\hat{a})+[1-\beta(t)]\underline{N}_i\underline{N}_i^T \tag{22}$$

after an appropriate delay of the Viterbi trellis, a decision is made that the most likely estimate for the sequence of symbols ... Here L is the noise correlation length and K=K₁+K₂+1 is the ISI length. Let the current estimate for the (L+1)x(L+1) covariance

The following example illustrates how the algorithm in (22) works. Assume a PR4 target response with a simple trellis structure as shown in FIG. 4 Notice that for PR4, the symbols can be equated to the trellis states, as is illustrated in FIG. 4. The number next to each branch in FIG. 4

The one-dimensional equivalent of equation (22) is

$$\hat{\sigma}_{new}{}^2=\beta\hat{\sigma}_{old}{}^2+[1-\beta]N_i^2. \tag{23}$$

This equation can be used in conjunction with the metric in (10).

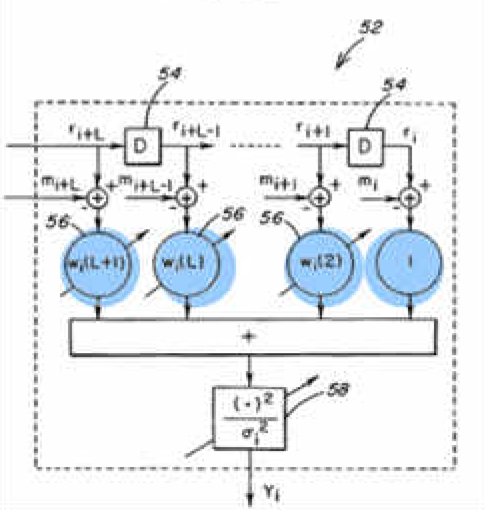
close to 1, while it should drop as the nonstationarity level increases. The forgetting factor is typically taken time-dependent to account for the start-up conditions of the RLS algorithm in (22). As more data is processed, a steady-state

The rank-one sample covariance matrix $\underline{N}\,\underline{N}^T$ is used to find the covariance matrix update:

CMU's "All Embodiments" Argument Fails

- The third method (no covariance matrix) corresponds to claims that calculate a "weight w_i" (Group III Claims)
  - These claims do not use "noise covariance matrices"

FIG. 3B



20. A branch metric computation circuit for generating a branch weight for branches of a trellis for a Viterbi-like detector, wherein the detector is used in a system having Gaussian noise, comprising:

a logarithmic circuit having for each branch an input responsive to a branch address and an output;

a plurality of arithmetic circuits each having a first input responsive to a plurality of signal samples, a second input responsive to a plurality of target response values, and an output, wherein each of the arithmetic circuits corresponds to each of the branches;

a sum circuit having for each branch a first input responsive to said output of said logarithmic circuit, a second input responsive to said output of said arithmetic circuit, and an output.

'839 Patent Claims 20-21

Correlated Noise

Claim Term	CMU's Construction	Marvell's Construction
correlated noise '839 Patent Claims 2 and 5 '180 Patent Claim 1	noise with 'correlation' among 'signal samples,' such as that caused by coloring by front-end equalizers, media noise, media nonlinearities, and magnetoresistive (MR) head nonlinearities. CMU Brf. at 19	noise having nonzero 'covariance' (see construction of 'covariance' above). Marvell Brf. at 32-33

- The Dispute
  - Should "correlated noise" be accorded its ordinary meaning in engineering and statistics (Marvell) or its lay meaning with a list of examples (CMU)?

MARVELL's PROPOSED CONSTRUCTION

"Correlated noise" means "noise having nonzero 'covariance' (see construction of 'covariance' below)."

3. Covariance

This correlation value is hard to place in context, so there is a related statistic, called covariance, which measures both the degree and the direction of the relationship between two sets of data. *See* Proakis Decl. at ¶ 22; P. Olofsson, Probability, Statistics and Stochastic Processes, at 200-201 (2005) (Exh. 19). The covariance is zero if two sets of data are not related, positive if larger data values in one set correspond to larger values in the other set, and

Processes, at 200-201 (2005) (Exh. 19). The covariance is zero if two sets of data are not related, positive if larger data values in one set correspond to larger values in the other set, and

negative if larger data values in one set correspond to smaller data values in the other set. Olofsson at 201. The "stronger" the dependence between two sets of data, the larger the covariance. *Id.* As its name suggests, covariance is a combination of correlation and variance. Covariance is calculated by subtracting the mean for each data point to obtain deviations before the multiplication and then taking the average of the products. Pocket Dictionary of Statistics ("covariance") (Exh. 16); Ross (Exh. 17 at 119). In the test score example, the mean and deviations for Test #1 were calculated above. The mean of Test #2 is 83, and the deviations for both Test #1 and Test #2 are shown in Table 2 below. To calculate the covariance of these two

Specification

- Patents describe correlated noise using mathematical terms

Correlation-sensitive branch metric. In the most general case, the correlation length is L>0. The leading and trailing ISI lengths are K_l and K_t, respectively. The noise is now considered to be both correlated and signal-dependent. Joint Gaussian noise pdfs are assumed. This assumption is well justified in magnetic recording because the experimental evidence shows that the dominant media noise modes have Gaussian-like histograms. The conditional pdfs do not factor out in this general case, so the general form for the pdf is:

$$\frac{f(r_{i+1}, \ldots, r_{i+L} \mid a_{i-K_l}, \ldots, a_{i+L+K_t})}{f(r_i, r_{i+1}, \ldots, r_{i+L} \mid a_{i-K_l}, \ldots, a_{i+L+K_t})} = \sqrt{\frac{(2\pi)^{L+1} det\ C_i}{(2\pi)^L det\ c_i}} \frac{\exp[\underline{N}_i^T C_i^{-1} \underline{N}_i]}{\exp[\underline{n}_i^T c_i^{-1} \underline{n}_i]} \quad (11)$$

'839 Patent 6:36-52

MARVELL's PROPOSED CONSTRUCTION

"Correlated noise" means "noise having nonzero 'covariance' (see construction of 'covariance' below)."

Signal-Dependent Noise

Claim Term	CMU's Construction	Marvell's Construction
signal-dependent noise '839 Patent Claims 2 and 5 '180 Patent Claim 1	media noise in the readback signal whose noise structure is attributable to a specific sequence of symbols (e.g., written symbols). CMU Brf. at 32	noise that is dependent on the signal. Marvell Brf. at 34-35

- Dispute:
  - Does "signal-dependent noise" have its ordinary meaning (Marvell) or should it be limited to a particular type of noise (media noise) found in magnetic recording (CMU)?